Accounting for Causality When Measuring Sales Lift from Television Advertising

Television Campaigns Are Shown To Be More Effective for Lighter Brand Users

HENRY ASSAEL
Stern School of Business, New York University
hassael@stern.nyu.edu

MASAKAZU ISHIHARA
Stern School of Business, New York University
mishihar@stern.nyu.edu

BAEK JUNG KIM
Sauder School of Business, University of British Columbia
baekjung.kim@sauder.ubc.ca

Establishing causality has long been an issue in the advertising-to-sales relationship. The ideal approach in establishing causality is randomized assignment of matched respondents to an exposure and non-exposure group. Given the difficulty in implementing such an experimental approach, observational data has been used in establishing a linkage between advertising and sales. The problem here is that the lack of randomized trials requires controlling for endogeneity (*i.e.,* factors other than advertising related to sales), and accounting for heterogeneity *(i.e,* variations in consumer characteristics affecting responses to advertising). The authors apply two state-of-the-art methods to single-source data so as to control for endogeneity and account for heterogeneity. Based on this analysis, the authors hope to improve on the use of observational data in accounting for advertising effects in a causal fashion.

INTRODUCTION

Researchers have relied on two types of data to establish causality in relating television advertising to sales: experimentation and observation. This paper focuses first on the limits of experimentation and considers utilizing observational data to establish causality by matching exposed and unexposed groups. A key consideration in estimating the causal effect of advertisements on sales by using observational data rather than experimentation is the need to control for endogeneity. Controlling for endogeneity requires accounting for variables correlated to sales independent of advertising effects. Along with controlling for endogeneity, substantial variation in how particular consumers respond to advertisements—namely, heterogeneous advertising effects—must also be accounted for. Estimating such effects requires accounting for differences in consumer characteristics that may condition advertising exposure independent of the influence of advertising on sales.

Management Slant

- Refined methods of statistical matching are an effective substitute for experimentation in accounting for causality in the television advertising-to-sales relationship.
- Such matching must account for both demographic characteristics and pre-campaign purchase behavior.
- In controlling for endogeneity and accounting for heterogeneity in the television advertising-to-sales relationship, causal results showed that the campaign for the focal brand was most effective in influencing light as opposed to heavy users.
- Failure to control for endogeneity results in the opposite effect; that is, the campaign is most effective in influencing heavy users.

Submitted May 1, 2019;
revised November 20, 2019;
accepted January 28, 2020;
published online November 2, 2020.

DOI: 10.2501/JAR-2020-024

In the report with an approach closest to that of the authors, Berkovich and Wood (2016) studied the direct effect of an advertising campaign on sales on the basis of a matched sample of advertising exposure and nonexposure households. They exploited demographics and purchase statistics in a matching process and compared purchase behaviors between those groups to measure sales lift. In the current paper, the authors significantly advance this approach by applying two state-of-the-art methods to control for endogeneity and account for heterogeneity in estimating the causal effects of television advertising on sales lift for a focal brand. Through better matching combined with better controls for endogeneity, the authors hope to improve the use of observational data in accounting for causality in the advertising-to-sales relationship. In so doing, the authors provide detailed procedures so that this paper helps expand the toolbox for practitioners.

METHODOLOGY

Experimentation versus Observational Data

Experimentation is the gold standard in measuring the causal effects of advertising on sales by randomly assigning respondents to an exposure group and a nonexposure group, thus ensuring that the response (sales) is a function of the experimental stimulus (advertising; *e.g.*, Aaker and Carman, 1982; Hu, Lodish, and Krieger, 2007; Lodish, Abraham, Kalmenson, Livelsberger, *et al.*, 1995). Experimentation has been used previously to test pre-campaign advertising effects through test markets (Eskin, 1975; Eskin and Baron, 1977). Experimentation is difficult to implement once a campaign is launched, because of the difficulty in ensuring the integrity of a control (nonexposure) group. Marketers using traditional media such as television and radio would find it difficult to ensure that a control group was not exposed to any advertising during the campaign period. They could utilize split-cable television, but directing resources is more likely to be a function of marketing rather than experimental objectives; namely, targeting the most likely buyers rather than assigning respondents to randomized groups. Marketers may be reluctant to withhold advertising from a randomized control group because of the sacrifice in revenue as a result. If test groups are selected by targeting criteria rather than randomization, this creates the same problems of endogeneity as for observational data; namely, that the relationship between advertising and sales could be artifactual in that sales results are independent of advertising exposure.

The primary alternative to experimentation is to use observational data (*e.g.*, single-source data relating advertising exposure to purchases at the individual level) and match those not exposed with those exposed on a variety of criteria (*e.g.*, Bronnenberg, Dubé, and Mela, 2010; Draganska, Hartmann, and Stanglein, 2014; Kumar, Bezawada, Rishika, and Janakiraman, *et al.*, 2016). The problem here is to ensure that matching treatment and control groups on key criteria is sufficiently randomized to control for endogeneity. This problem has been reported in previous research (Blake, Nosko, and Tadelis, 2015; Lewis, Rao, and Riley, 2011).

Measuring Sales Lift

The authors define sales lift as the difference in sales results before and immediately after an advertising campaign resulting from the campaign. Berkovich and Wood (2016) explained that, to be included in a sales lift study, households must meet pre- and post-campaign periods “static” to ensure that they are actively purchasing the focal product. This static is a function of various shopping behaviors, such as purchase frequency, purchase spending, and interpurchase time. Similarly, the current authors measure sales lift by using different household samples as a function of purchase frequency and purchase spending on the focal brand and category during the pre- and post-campaign periods (*e.g.*, top 10 percent, 25 percent, and 50 percent of total households in terms of purchase frequency and spending for the focal brand). In this study, the pre-campaign period is defined as one year before the start of the television campaign. Although Berkovich and Wood (2016) measured household-level incremental sales considering matched samples between test and control groups, they primarily relied on demographics and purchase statistics in a matching process and then simply compared statistics of purchase behaviors between those groups to measure sales lift. In contrast, this study exploits formal econometric approaches—propensity score matching and difference in differences—for matching samples and measuring sales lift, and will discuss all details of the empirical methods in the subsequent section.

Data and Summary Statistics

This study tests sales lift for a brand in a chocolate-candy category for a television campaign that ran from March 28 to June 16, 2016. Sales data refer to household-based scanner data collected by Nielsen Catalina Solutions for a sample of households in the United States whose purchases are scanned and recorded after each shopping trip. The data are single source also, combining television advertising exposure and purchase data from the same household at the individual level so that the effects of television advertising on purchase behaviors can be studied. There are data for 435,327 households in total, including all shopping information from March 29, 2015, to June 26, 2016. (See Table 1 for the summary statistics for all the details of purchase behavior in the chocolate-candy product category.)

Table 1 Summary Statistics: Total Samples and Total Periods, Pre- and Post-campaign

Variable	Mean	SD	Top 75%	Top 50% (Median)	Top 25%	Top 10%	Top 1%
Category purchase frequency	10.13	11.32	≥3.00	≥7.00	≥13.00	≥22.00	≥54.00
Category purchase net amount ($ per trip)	4.51	3.11	≥2.59	≥3.89	≥5.58	≥7.81	≥15.16
Category purchase gross amount ($ per trip)	5.16	3.53	≥3.00	≥4.46	≥6.40	≥8.90	≥17.28
Brand purchase frequency	1.64	3.52	≥0.00	≥1.00	≥2.00	≥4.00	≥7.00
Brand purchase net amount ($ per trip)	3.23	2.45	≥1.59	≥2.72	≥4.00	≥5.99	≥11.99
Brand purchase gross amount ($ per trip)	3.74	2.76	≥1.92	≥3.23	≥4.64	≥6.96	≥13.35

Controlling for Endogeneity

Because this study uses observational data, controlling for endogeneity is central to the analysis. Two sources of endogeneity in the relationship between advertising and sales are targeting-induced endogeneity and exposure-induced endogeneity (Gordon, Zettelmeyer, Bhargava, and Chapsky, 2019; Lewis *et al*., 2011). Targeting-induced endogeneity occurs when marketers target heavy buyers, resulting in purchases that would have occurred independent of advertising effects. Relating advertising to sales overstates the association. Exposure-induced endogeneity is a function of variations in advertising exposure independent of advertising effects. If demographic characteristics are related to the likelihood of viewing a campaign, exposure is likely to be a function of such characteristics as well as advertising effectiveness. Relating advertising effectiveness to sales must control for demographic covariates related to the level of advertising exposure.

Additionally, previous literature (*e.g*., Anand and Shachar, 2011; Lovett and Peress, 2015) particularly consider endogeneity of television advertising that is mainly driven by unobserved characteristics such as geography or demographics. The most recent study investigated the problem of individual-level television advertisement targeting by leveraging granular data linking household-level television advertisement viewing with product purchase (Tuchman, Nair, and Gardete, 2018). Although these papers rigorously control for advertising endogeneity using a structural model, it would be difficult for practitioners to implement and replicate this strategy because of a sophisticated analysis process. Therefore, the authors propose another approach to control for television advertising endogeneity in a reduced-form way that is simpler and easier to execute.

The authors combine two state-of-the art econometric approaches, difference in differences and propensity score matching, to measure television advertising effects after controlling for the endogeneity of television advertising exposure. Ever since the work of Ashenfelter and Card (1985), the use of the difference in differences method has become widespread in marketing and economics (*e.g*., Angrist and Pischke, 2009; Goldfarb and Tucker, 2014; Imbens and Wooldridge, 2009; Meyer, 1995). In this approach, outcomes (*i.e*., purchases of the focal brand) are observed for two groups (the advertisement-exposed group and the non-advertisement-exposed group) for two time periods (pre- and post-campaign periods).

Specifically, the following regression equation was used:

$$\begin{aligned}\text{Purchase}_{it} = \beta_0 + \beta_1 \cdot \text{Ads Exposure}_i + \beta_2 \cdot \text{Campaign Period}_t \\ + \beta_3 \cdot \text{Ads Exposure}_i \cdot \text{Campaign Period}_t \\ + \beta_4 \cdot \text{Price}_{it} + \varepsilon_{it}\end{aligned} \tag{1}$$

where *i* indicates a household and *t* indicates a shopping occasion. Purchase_{it} is an indicator equal to 1 if household *i* purchased the focal brand at time *t*. Similarly, Ad Exposure_i is an indicator equal to 1 if household *i* was exposed to a television advertisement at least once during the campaign period and 0 if otherwise. Campaign Period_t is also an indicator equal to 1 if time *t* is after the starting date of the television advertising campaign; otherwise, it is equal to 0. The final variable is $\text{Price}_{it.}$ Last, ε_{it} is a stochastic error term.

First, price is one of the most important components affecting consumer purchase in consumer-packaged goods, and it is expected to influence purchase response negatively. Second, to identify the effect of television advertisements on purchase, Ad Exposure*i*, Campaign Period*t*, and the interaction term must be measured. One of the groups is exposed to television advertisements in the campaign period but not in the pre-campaign period. The second group is not exposed to television advertisements during either period. In the case in which the same units within a group are observed in each time period, the average gain in the second (control) group can be subtracted from the average gain in the first (treatment) group. This removes biases in the second period comparisons between the treatment and control groups that could be the results of permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the results of trends.

As for the interpretation of the coefficients, β_1 captures possible differences between the television-advertising-exposed and non-advertising-exposed groups. The coefficient of the time period, dummy β_2, captures aggregate factors that would affect changes in the purchase of our focal brand, even in the absence of the television advertising campaign. The authors aim to determine coefficient β_3, which accounts for the interaction term between treatment status and behavior in the pretreatment period.

$$\beta_3 = \left(\text{purchase}_{campaign,ad\ exposure} - \text{purchase}_{precampaign,ad\ exposure}\right) \left(\text{purchase}_{campaign,no\ ad\ exposure} - \text{purchase}_{precampaign,no\ ad\ exposure}\right) \quad (2)$$

β_3 measures the incremental advertising effects of those exposed to the campaign versus those not exposed while accounting for pre-campaign behavior.

This coefficient measures the causal effects of television advertising exposure during the campaign period on brand purchases. The authors hypothesize that β_3 is positive and significant if television advertisements have positive causal effects on brand purchases. The authors run the ordinary least squares regression of the equation, a linear probability model.

As mentioned previously, however, the advertising exposure variable in Equation 1 is endogenous because of the exposure selection issue. To control for the endogeneity of television advertising exposure, this paper uses another state-of-the-art econometric technique: propensity score matching. This approach is used in many academic works in marketing; for example, examining the effectiveness of Internet and television advertisements on brand building (Draganska *et al*., 2014), understanding the effects of firm-generated social media on customer metrics (Kumar *et al*., 2016), and investigating the effects of access to digital video recorders on sales of advertised products (Bronnenberg *et al*., 2010). The propensity score is the probability of treatment assignment (*i.e*., status of being exposed to television advertisements) conditional on observed baseline characteristics (*i.e*., demographics and precampaign period purchase behaviors). The propensity score allows the design and analysis of an observational (nonrandomized) study so that it mimics some of the particular characteristics of a randomized controlled trial. In particular, the propensity score is a balancing score so, conditional on the propensity score, the distribution of the observed baseline covariates will be similar for advertising-exposed and non-advertising-exposed households.

To compute the propensity score, the authors used the following equation:

$$\Pr\left(\text{Ads Exposure}_i = 1 \middle| X_i\right) = \frac{\exp\left(X_i\right)}{1+\exp\left(X_i\right)}, \quad (3)$$

where i indicates a household and Ad Exposurei is an indicator equal to 1 if household i was exposed to television advertisements at least once during the campaign period and 0 otherwise. The important term is X_i, which is a vector of covariates, demographics, and pre-campaign period purchase behaviors (including income, age, household size, race, gender, population of the household's location, marital status, education, origin, purchase spending, purchase amount, purchase with promotion at both the brand and category levels, and main shopping retailers).

On the basis of the propensity scores, the nearest neighbor is matched, enabling matching advertising-exposed households with non-advertising-exposed households so that the two groups can be compared in a process similar to that of a randomized control trial.

The authors found a close match between households exposed and those not exposed to advertising in demographic characteristics and a lack of statistically significant difference in pre-campaign purchase behavior as covariates, which might affect the probability of being exposed to advertisements (See Table 2). A balanced distribution of the covariates affecting advertising exposure probability between these two groups shows how we control for the endogeneity of television advertising exposure.

This approach is similar to the one used by Berkovich and Wood (2016) to match test and control households on the basis of household demographics and shopping characteristics. They stated that the purpose of matching is to be able to attribute any increase in spending to advertising rather than to an existing preference for the product being studied. The current study goes further by matching the households on the basis of their propensity scores of exposure to television advertisements, which is a function of demographics and pre-campaign period purchase behaviors. Specifically, the current study computes the propensity scores of all households in the data and chooses one of the households in the control group (*i.e*., a household not exposed to television advertising) that shows the nearest propensity score of the household in the treatment group (*i.e*., a household exposed to television advertising).

After the matching process, there is now a dummy variable, advertising exposure, which is equal to 1 if household i is exposed to advertisements and 0 otherwise for all households. This newly created advertising exposure variable is used as a covariate in Equation 1. By doing so, the effects of the television advertising campaign on brand purchase after controlling for the endogeneity of advertising exposure can be measured (See Table 4, shown later).

Heterogeneous Advertising Effects

The estimates in previous sections (*e.g*., β_3 in Equation 1) indicate average causal effects of advertising on brand purchase, implying that all the households exposed to the advertising campaign have the same sensitivity to television advertisements. How the effects

Table 2 Share of Households in Groups Exposed and Not Exposed to Advertising

Category and Variable	Share, in %, of HH (and SD): Advertising Exposure	Share, in %, of HH (and SD): No Advertising Exposure
Income ($)		
0	1.39 (.117)	1.43 (.120)
1–10,000	2.63 (.160)	2.68 (.163)
10,001–15,000	4.71 (.212)	4.69 (.216)
15,001–20,000	6.45 (.246)	6.43 (.249)
20,001–30,000	7.88 (.269)	7.89 (.275)
30,001–40,000	5.14 (.221)	5.10 (.225)
40,001–50,000	12.05 (.326)	12.09 (.322)
50,001–60,000	12.24 (.328)	12.20 (.329)
60,001–75,000	20.45 (.403)	20.42 (.397)
75,001–100,000	10.37 (.305)	10.45 (.302)
100,001–125,000	16.66 (.373)	16.61 (.374)
Age		
less than 25 years	1.72 (.130)	1.68 (.127)
25–34 years	6.16 (.240)	6.22 (.238)
35–44 years	12.43 (.330)	12.41 (.319)
45–54 years	22.74 (.419)	22.66 (.410)
55–64 years	27.76 (.448)	27.84 (.447)
65 years and more	29.18 (.455)	29.18 (.466)
Race		
Asian	1.46 (.120)	1.49 (.121)
Black	7.55 (.264)	7.47 (.263)
Hispanic	4.99 (.218)	5.01 (.218)
White	85.70 (.350)	85.74 (.350)
Unknown	0.30 (.055)	0.29 (.054)

Category and Variable	Share, in %, of HH (and SD): Advertising Exposure	Share, in %, of HH (and SD): No Advertising Exposure
Gender		
Female	16.85 (.374)	16.97 (.379)
Male	83.15 (.374)	83.03 (.379)
Population (U.S. HH in metro area)		
40%	30.00 (.458)	30.00 (.457)
30%	32.74 (.469)	32.71 (.467)
15%	24.50 (.430)	24.48 (.435)
All others	12.75 (.334)	12.81 (.332)
Marital status		
Married	81.96 (.385)	82.00 (.392)
Single	8.09 (.273)	8.02 (.274)
Missing	8.08 (.272)	8.15 (.286)
Unknown	1.87 (.136)	1.83 (.132)
Education		
Elementary school	7.92 (.270)	7.92 (.273)
Middle school	28.97 (.454)	29.04 (.455)
High school	27.05 (.444)	27.07 (.440)
College	36.07 (.480)	35.96 (.482)
Average brand		
Spending amount	3.824 (8.237)	3.842 (8.493)
Promotion amount	0.566 (1.551)	0.580 (1.606)
Purchase quantity	1.756 (3.670)	1.781 (3.930)
Average category:		
Spending amount	38.064 (42.623)	37.101 (40.570)
Promotion amount	5.288 (7.769)	5.346 (8.025)
Purchase quantity	16.584 (18.283)	16.576 (18.923)

Note: HH = household, SD = standard deviation.

vary across subhouseholds is also of interest, since different subgroups of the households are likely to react differently to television advertisements. Previous literature in marketing and economics (*e.g.*, Berkovich and Wood, 2016; Rossi and Allenby, 1993; Zenetti and Otter, 2014) emphasizes the importance of such heterogeneous advertising effects on consumer purchase decisions.

To capture heterogeneous advertising effects, this study exploits a more robust analysis that is an extension of the difference in differences analysis described earlier: a difference in difference in differences analysis using a different control group within the treatment state (*i.e.*, advertisement exposure). The causal effects of television advertisements (*i.e.*, β_3 in Equation 1) in the difference in differences analysis imply that all households that were exposed to advertisements have the same sensitivity in responding to advertisements. The effectiveness of the advertisement, however, could vary across households, depending on their purchase behaviors. Heavy brand purchasers, for example, may be more (or less) likely to respond to advertisements compared with light brand purchasers. To understand heterogeneous causal advertising effects on brand purchase, more robust indicator variables are needed: the top 50 percent, top 25 percent, and top 10 percent of brand purchasers based on the total purchase frequency before the campaign period. The indicator variable for the top 50 percent is equal to 1 if the household purchased the study's focal brand at least twice in the pre-campaign period and 0 if otherwise. The authors create the other indicators in the same manner; for example, the top 25 percent

Table 3 Number of Households that Purchased Brand during the Pre-campaign Period

Percentage of Category Purchasers	Number (and %) of Households
Top 50% (brand purchase at least twice)	97,995 (34.5)
Top 25% (brand purchase at least three times)	65,903 (23.2)
Top 10% (brand purchase at least six times)	25,311 (8.9)

Table 4 Causal Effects of Advertising on Brand Purchase Controlling for Endogeneity

	Dependent Variable: Brand Choice	
Variable	**Nonmatching**	**Matching**
Advertising Exposure × Campaign Period (β_3) (1)	–0.002 (0.002)	**0.005***** **(0.002)**
Log (price)	–0.004*** (0.001)	–0.008*** (0.001)
Fixed effects		
Individual	Yes	Yes
Shopping occasion	Yes	Yes
Observations	4,410,322	2,535,493
R^2	0.258	0.265

Note: Standard errors (in parentheses) are clustered at the consumer level.
*p < 0.1; **p < 0.05; ***p < 0.01.

of brand purchasers equals 1 if they bought the brand at least three times in the pre-campaign period and 0 if otherwise. (See Table 3 for a summary of the definition, total number, and actual percentage of these household segments.)

Next, several interaction terms are added in Equation 1 as follows, using the top 50 percent as an example:

$$\begin{aligned}\text{Purchase}_{it} = {} & \beta_0 + \beta_1 \cdot \text{Ads Exposure}_i + \beta_2 \cdot \text{Campaign Period}_{it} \\ & + \beta_3 \cdot \text{Top50\%} + \beta_4 \cdot \text{Ads Exposure}_i \cdot \text{Campaign Period}_{it} \\ & + \beta_5 \cdot \text{Ads Exposure}_i \cdot \text{Top50\%} \\ & + \beta_6 \cdot \text{Campaign Period}_{it} \cdot \text{Top50\%} \\ & + \beta_7 \cdot \text{Ads Exposure}_i \cdot \text{Campaign Period}_{it} \cdot \text{Top50\%} \\ & + \beta_8 \cdot \text{Price}_{it} + \varepsilon_{it}\end{aligned} \tag{4}$$

As for the interpretation of the coefficients, β_4 captures the causal effects of the campaign on households that have never purchased or purchased at least once the focal brand in the pre-campaign period (*i.e.*, households that are not in the top 50 percent). β_7 denotes the difference in causal effects of the advertising campaign between households that were exposed to the campaign and purchased the focal brand at least twice in the pre-campaign period versus those that did not purchase or purchased at least once (β_4) while controlling for prepurchase behavior. Similarly, replacing the Top50% with the other indicator variables, such as Top25% and Top10%, allows for the heterogeneity of advertising effects between corresponding groups to be calculated. This approach is called the difference in differences in differences. Here, β_7 is described as follows:

$$\begin{aligned}\beta_7 = {} & \left(\text{purchase}_{campaign,\, top50\%,\, ad\ exp} - \text{purchase}_{precampaign,\, top50\%,\, ad\ exp}\right) \\ & -\left(\text{purchase}_{campaign,\, NONtop50\%,\, ad\ exp} - \text{purchase}_{precampaign,\, NONtop50\%,\, no\ ad\ exp}\right) \\ & -\left(\text{purchase}_{campaign,\, top50\%,\, no\ ad\ exp} - \text{purchase}_{precampaign,\, top50\%,\, no\ ad\ exp}\right)\end{aligned} \tag{5}$$

Again, β_7 measures the difference in causal effects of television advertisements between the top 50 percent of households (purchased at least twice) and everyone else.

RESULTS

The authors report the causal effects of the television advertising campaign on brand purchase, controlling for endogeneity (See Table 4). The second column of Table 4 shows the results from the difference in differences analysis without a matching process to control for advertising exposure. The third column shows the results from the difference in differences analysis using both demographics and pre-campaign period purchase behaviors in the matching process.

The key measure is β_3, the interaction between advertising exposure and the campaign period. In the second column, after matching with consumer demographics and pre-campaign purchase behavior, the marginal effect of the interaction between advertising exposure and campaign period is 0.005 (*i.e.*, the incremental effects of β_3 in Equation 1), meaning that the households that were exposed to advertisements during the campaign period more likely will purchase the brand by 0.5 percentage points, compared with the households that were not exposed to advertisements during the campaign period. Given that the average brand purchase probability across all observations is 15.17 percentage points, the 0.5 percentage point increase corresponds to a 3.30 percent increase in brand purchase probability (0.5/15.17).

The interesting observation is that if advertising exposure endogeneity is not controlled, there is little effect of advertising exposure on brand purchase. Without a matching process, the coefficient of the interaction term becomes insignificant (See Table 4, column 1). That means households that were exposed to advertisements during the campaign period are statistically indifferent from the households that were not exposed to advertisements during the campaign period. The findings demonstrate the importance of accounting for pre-campaign purchases and demographic matching to show advertising effects.

Table 5 Heterogeneous Advertising Effects by Brand Purchase Frequency, with Demographics and Pre-campaign Purchase Matching

Variable	Dependent Variable: Brand Choice for Brand Purchase HH Segment		
	Top 50% vs. Bottom 50%	**Top 25% vs. Bottom 75%**	**Top 10% vs. Bottom 90%**
Advertising Exposure × Campaign Period (β_4)	**0.008***** **(0.002)**	**0.006***** **(0.002)**	**0.004***** **(0.002)**
Brand Purchase HH Segment × Campaign Period	–0.107*** (0.003)	–0.099*** (0.004)	–0.089*** (0.006)
Advertising Exposure × HH Segment × Campaign Period (β_7)	**–0.018***** **(0.004)**	**–0.017***** **(0.004)**	**–0.013***** **(0.006)**
Log (price)	–0.008*** (0.001)	–0.008*** (0.001)	–0.008*** (0.001)
Fixed effects			
Individual	Yes	Yes	Yes
Shopping occasion	Yes	Yes	Yes
Observations	2,535,493	2,535,493	2,535,493
R^2	0.268	0.267	0.266

Note: All standard errors (shown in parentheses) are clustered at the consumer level. HH = household. *p < 0.1; **p < 0.05; ***p < 0.01.

The authors report the heterogeneous causal effects of television advertisements from the estimations of the difference in difference in differences analysis (See Table 5). Here, heterogeneity is accounted for from brand purchase frequency in the pre-campaign period. As the purpose of allowing for heterogeneity in this analysis is to distinguish the brand choice between heavier and lighter brand users, it can be observed how television advertisements influence consumers by frequency of brand purchase. Each column shows results of different segments, with brand purchases defined at the top 50, 25, and 10 percentiles (See Table 3). The authors use both demographics and pre-campaign period purchase behaviors in the matching process to control for the endogeneity of television advertisement exposure.

The results show when households are split using the indicator variable for the top 50 percent versus bottom 50 percent of brand purchasers; that is, households that purchased the focal brand at least twice in the pre-campaign period (brand purchasers) versus those that did not or purchased at least once (See Table 5, column 2). β_4 measures the causal incremental effect of television advertisements for nonpurchasers or lighter brand purchasers. β_7 measures the incremental difference in causal effect of the campaign for heavier brand purchasers versus nonpurchasers or lighter brand purchasers in the pre-campaign period. The results show a significant effect of the campaign on lighter brand purchasers or nonpurchasers, with an incremental increase in brand purchases of 0.8 percentage points; the campaign, however, had a smaller effect on heavier brand purchasers. β_7 shows a decrease of an additional 1.8 percentage points in brand purchases for heavier brand purchasers. The results suggest that the advertising campaign increases brand purchase for both purchasers and nonpurchasers but more so for lighter brand purchasers and nonpurchasers.

The results also show when households are split using the top 25 percent indicator variable; that is, households that purchased at least three times in the pre-campaign period versus those that did not purchase or who purchased, at most, twice—the bottom 75 percent (See Table 5, column 3). The results show that the causal effect of television advertisements for those in the bottom 75 percent (β_4) is a significant incremental increase of 0.6 percentage points in brand purchases. The difference in the incremental causal effect between households in the top 25 percent and those in the bottom 75 percent (β_7) is negative (and significant), and this result suggests that the causal effect of television advertisements for heavier brand purchasers is less effective than for nonbrand and lighter brand purchasers. A similar conclusion can be drawn from the results, where households are split using the top 10 percent indicator variable; those who purchased at least six times versus those who purchased, at most, five times or did not purchase (See Table 5, column 4).

In short, the results in Table 5 suggest that the campaign was effective across the board for both users and nonusers but most effective for nonusers and lighter users. Targeting heavy brand users would yield no advantage in increasing brand purchases.

Next, the authors further investigate how the results of heterogeneous advertisement effect controlling for endogeneity differ from the results of not controlling for it. The authors conduct the same analysis with the sample without matching (See Table 6). Similar to the results in the previous table, each column shows results of different segments, with brand purchases defined at the top 50, 25, and 10 percentiles of all the households in the sample (*i.e.*, 435,327). Interestingly, the coefficients of the three-way interaction terms (*i.e.*, β_7) in all the segments are positive and significant, implying that advertising effects are greater for heavier brand purchasers compared with nonpurchasers or lighter brand purchasers. Indeed, this result drives an opposite insight to the findings in Table 5 and thus shows why controlling for advertising endogeneity is important for understanding the correct advertisement effect on brand purchase.

Overall, the findings suggest that, for this chocolate-candy brand at least, a broader based targeting to light brand users and nonusers

Table 6 Heterogeneous Advertising Effects by Brand Purchase Frequency, without Matching

Variable	Dependent Variable: Brand Choice for Brand Purchase HH Segment		
	Top 50% vs. Bottom 50%	**Top 25% vs. Bottom 75%**	**Top 10% vs. Bottom 90%**
Ad Exposure × Campaign Period (β_4)	**0.006***** **(0.001)**	**0.008***** **(0.001)**	**0.006***** **(0.001)**
Brand Purchase HH Segment × Campaign Period	–0.113*** (0.003)	–0.101*** (0.004)	–0.087*** (0.006)
Ad Exposure × HH Segment × Campaign Period (β_7)	**0.014***** **(0.003)**	**0.012***** **(0.004)**	**0.010*** **(0.006)**
Log (price)	–0.004*** (0.001)	–0.004*** (0.001)	–0.004*** (0.001)
Fixed effects			
Individual	Yes	Yes	Yes
Shopping occasion	Yes	Yes	Yes
Observations	4,410,322	4,410,322	4,410,322
R^2	0.261	0.260	0.259

Note: All standard errors (in parentheses) are clustered at the consumer level.
HH = household. *p < 0.1; **p < 0.05; ***p < 0.01.

is more effective than a more concentrated targeting to the heavy quarter or heavy tenth of brand users.

DISCUSSION

Limitations and Future Research

The data are derived from a single brand source. There is no information on switching behavior and, as a result, the authors cannot analyze competitive influences on the advertising-to-sales relationship in a dynamic marketplace. This means that, although it is possible to measure causal advertising effects on sales for the focal brand, brand managers do not know where a sales increment came from as a result of the advertising campaign. Future research could build on this analysis by investigating a multibrand data source that could show brand managers how consumers substitute their brand purchases due to an advertising campaign. Examining the causal effects that advertisements have on sales by considering a substitution pattern between brands would provide better insight into the product category. Moreover, this analysis focused on one category. Future research could focus on causal advertising effects across categories.

Conclusion

Controlling for advertising endogeneity and accounting for heterogeneity in the campaign for the focal brand, this analysis confirms several findings. First and most important, without matching for pre-campaign purchases, advertising effects in the campaign period become insignificant. This demonstrates the importance of matching by purchase predispositions to account for advertising effects in evaluating the advertising-to-sales relationship. Second, the campaign was effective in influencing both brand purchasers and nonbrand purchasers but more effective for nonbrand purchasers and lighter brand purchasers compared with heavy brand purchasers. Third, without a matching process to control for endogeneity, the analysis on heterogeneous advertisement effects shows incorrect insight (*i.e.*, advertising to heavier brand purchasers is more effective). This demonstrates why using the proposed approaches is important.

The analytical approach illustrated for controlling for endogeneity in accounting for the causal effects of advertising and showing its heterogeneous effects is, perhaps, more significant. JAR

ACKNOWLEDGMENTS

The authors thank Leslie Wood, chief research officer, and Andrew Faehnle, director of research and development, at Nielsen Catalina Solutions for the data used in the analysis and for their guidance in establishing the analytical approach.

ABOUT THE AUTHORS

Henry Assael is Professor Emeritus of Marketing at the Stern School of Business, New York University. He has published twelve articles in this journal over the years in the area of evaluating advertising and media effectiveness, most notably his article titled "From Silos to Synergy" on cross-media effectiveness that won the inaugural *JAR* Best Paper award in in 2011. Assael has authored the following texts: *Consumer Behavior and Marketing Action* (seven editions); *Marketing* (three editions); and *Marketing: Principles and Strategy*.

Masakazu Ishihara is associate professor of marketing at Stern School of Business, New York University. His research focuses on quantitative marketing and empirical industrial organization with particular interest in the dynamic effects of marketing strategies, forward-looking decision making by consumers and firms, and marketing and economics in the entertainment industry. Ishihara received the inaugural 2010 INFORMS Society for Marketing Science (ISMS) Doctoral Dissertation Competition Award, and Honorable Mention of Dick Wittink Prize Award in 2011. He has also been selected as one of the Marketing Science Institute (MSI) Young Scholars of 2017.

Baek Jung Kim is assistant professor of marketing at Sauder School of Business, University of British Columbia. He coauthored this paper as a doctoral student at NYU Stern School of Business. Kim's research interests include digital platform adoption and usage.

REFERENCES

Aaker, D., and J. M. Carman. "Are You Over-advertising?" *Journal of Advertising Research* 22, 4 (1982): 57–70.

Anand, B. N., and R. Shachar. "Advertising, the Matchmaker." *The RAND Journal of Economics* 42, 2 (2011): 205–245.

Angrist, J. D., and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press, 2008.

Ashenfelter, O., and D. Card. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *Review of Economics and Statistics* 67, 4 (1985): 648–660.

Berkovich, P., and L. Wood. "Using Single-Source Data to Measure Advertising Effectiveness." *Nielsen Journal of Measurement* 1, 2 (2016). Retrieved from the Nielsen website: https://www.nielsen.com/us/en/insights/report/2016/using-single-source-data-to-measure-advertising-effectiveness/

Blake, T., C. Nosko, and S. Tadelis. "Consumer Heterogeneity and Paid Search Effectiveness: A Large Scale Field Experiment." *Econometrica* 83, 1 (2015): 155–174.

Bronnenberg, B. J., J.-P. Dubé, and C. F. Mela. "Do Digital Video Recorders Influence Sales?" *Journal of Marketing Research* 47, 6 (2010): 998–1010.

Draganska, M., W. R. Hartmann, and G. Stanglein. "Internet versus Television Advertising: A Brand-Building Comparison." *Journal of Marketing Research* 51, 5 (2014): 578–590.

Eskin, G. J. "A Case for Test Market Experiments." *Journal of Advertising Research* 15, 2 (1975): 27–34.

Eskin, G. J., and P. H. Baron. "Effects of Price and Advertising in Test-Market Experiments." *Journal of Marketing Research* 14, 4 (1977): 499–508.

Goldfarb, A., and C. E. Tucker. (Working paper). "Conducting Research with Quasi-Experiments: A Guide for Marketers." Toronto, ON, Canada: Rotman School of Management, 2014.

Gordon, B. R., F. Zettelmeyer, N. Bhargava, and D. Chapsky. "A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook." *Marketing Science* 38, 2 (2019): 193–225.

Hu, Y., L. M. Lodish, and A. M. Krieger. "An Analysis of Real World TV Advertising Tests: A 15-Year Update." *Journal of Advertising Research* 47, 3 (2007): 341–353.

Imbens, G. W., and J. M. Wooldridge. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47, 1 (2009): 5–86.

Kumar, A., R. Bezawada, R. Rishika, R. Janakiraman, et al. "From Social to Sale: The Effects of Firm-Generated Content in Social Media on Customer Behavior." *Journal of Marketing* 80, 1 (2016): 7–25.

Lewis, R., J. Rao, and D. Riley. "Here, There, and Everywhere: Correlated Online Behaviors Can Lead to Overestimates of the Effects of Advertising." In *WWW '11: Proceedings of the 20th International Conference on World Wide Web*. New York: Association for Computing Machines, 2011.

Lodish, L. M., M. Abraham, S. Kalmenson, J. Livelsberger, et al. "How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Level T.V. Advertising Experiments." *Journal of Marketing Research* 32, 2 (1995): 125–139.

Lovett, M., and M. Peress. "Targeting Political Advertising on Television." *Quarterly Journal of Political Science* 10, 3 (2015): 391–432.

Meyer, B. D. "Natural and Quasi-Experiments in Economics." *Journal of Business & Economic Statistics* 13, 2 (1995): 151–161.

Rossi, P. E., and G. M. Allenby. "A Bayesian Approach to Estimating Household Parameters." *Journal of Marketing Research* 30, 2 (1993): 171–182.

Tuchman, A. E., H. S. Nair, and P. M. Gardete. "Television Ad-Skipping, Consumption Complementarities and the Consumer Demand for Advertising." *Quantitative Marketing and Economics* 16, 2 (2018): 111–174.

Zenetti, G., and T. Otter. "Bayesian Estimation of the Random Coefficients Logit from Aggregate Count Data." *Quantitative Marketing and Economics* 12, 1 (2014): 43–84.